



PUBLIC

Date: 17 May 2023

Author: UAB

Reviewed by: ISLEUK

Approved by: UAB

Version: V1.2



D8.3. DATA MANAGEMENT PLAN – INTERMEDIATE



The VIVALDI project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101000441.



History Chart

Version	Date	Implemented by
V1.2 (D8.3)	17 May 2023	UAB
V1.1 (D8.3)	02 May 2023	UAB
V1.0 (D8.2)	29 November 2021	UAB
Draft (D8.2)	15 November 2021	UAB

VIVALDI Consortium

Coordinator: UNIVERSITAT AUTÒNOMA DE BARCELONA (UAB)

Associated Beneficiaries:

1. UNIVERSITAT AUTONOMA DE BARCELONA (UAB)
2. UNIVERSITAET FUER BODENKULTUR WIEN (BOKU)
3. LULEA TEKNISKA UNIVERSITET (LTU)
4. VLAAMSE INSTELLING VOOR TECHNOLOGISCH ONDERZOEK N.V. (VITO)
5. HELMHOLTZ-ZENTRUM FUR MWELTFORSCHUNG GMBH (UFZ)
6. ACONDICIONAMIENTO TARRASENSE ASSOCIACION (LEITAT)
7. PROCESSION (PROC)
8. FUNDACIÓ UNIVERSITARIA BALMES (UVIC)
9. CO₂ VALUE EUROPE AISBL (CVE)
10. ISLE UTILITIES LIMITED (ISLEUK)
11. NUTRITION SCIENCES (NUT)
12. AVANTIUM CHEMICALS BV (AVT)
13. SUNPINE AB (SUNP)
14. COMPAÑIA CERVECERA DAMM, S.L (DAMM)
15. BIOAGRA SPOLKA AKCYJNA (BZK)
16. NOVAMONT SPA (NVMT)



No part of this document may be reproduced and/or published by print, photoprint, microfilm or any other means without the previous written consent of the VIVALDI consortium. The content of this deliverable does not reflect the official opinion of the European Union. Responsibility for the information and views expressed herein lies entirely with the author(s).



Executive Summary

This deliverable is the updated intermediate version of the Data Management Plan (DMP) that defines the data management policies that are followed during the execution of VIVALDI project and provides an overview of all the datasets that are collected and generated. It describes the types of data that are collected, processed or generated, how the data is handled during and after the project and the measures that are taken to ensure that the data collected or generated follows the FAIR principles of EU Horizon 2020 projects, ensuring that the data is findable, accessible, interoperable and reusable. The plan discusses the used metadata standards and identifiers, repositories for making the data available as well as quality assurance methods and reusability.

VIVALDI participates in the Horizon 2020 Open Research Data Pilot, meaning that the partners are committed to give open access to the research data generated, following the 'as open as possible, as closed as necessary' -principle.

The DMP is a living document, and it will be updated when significant changes are taking place that affect the type or handling of the data. Updated DMP will be again submitted at the end of the project (M48).



Table of Contents

1. Purpose of the data collection and generation	5
2. Collected and generated data.....	7
2.1. Reuse of existing data	7
2.2. Origin of the data.....	8
2.3. Size of the data.....	8
2.4. Utilisation.....	9
3. FAIR Data	10
3.1. Making data findable.....	10
3.2. Making data available.....	12
3.3. Making data interoperable	13
3.4. Increasing data reuse	14
4. Quality assurance.....	14
5. Costs	15
6. Data security.....	15
7. Ethical aspects	15
8. KPI Management Plan.....	16
9. Additional resources	16
10. Annexes	16
ANNEX I: Questionnaire regarding collected data	17
ANNEX II: CORA Metadata template	20



List of abbreviations

3-HP	3-hydroxypropionic acid
AB	Advisory Board
BIs	Bio-based Industries
DMP	Data Management Plan
DOI	Digital Object Identifier
EGRETA	Environment of Research and Transfer Management
EU	European Union
FA	Formic Acid
GA	Grant Agreement
GDPR	General Data Protection Regulation
GHG	Greenhouse Gas
IA	Itaconic Acid
KPI	Key Performance Indicator
LA	Lactic Acid
LCA	Life Cycle Analysis
LCC	Life Cycle Cost
MDC	Microbial Desalination Cell
MEC	Microbial Electrolysis Cell
MeOH	Methanol
SA	Succinic Acid
SOTA	State-of-the-art
USB	Universal Serial Bus
WP	Work Package
WPL	Work Package Leader



1. Purpose of the data collection and generation

VIVALDI is developing, validating and assessing the techno-economic and environmental feasibility of an innovative bio-based value-chain for the conversion of CO₂ emissions coming from bio-based industries (BIs) into selected added-value organic acids. Selectively captured CO₂ emissions are converted into feedstock for sustainable bioproduction of four organic acids with diverse market shares: 3-hydroxypropionic acid (3-HP), succinic acid (SA), itaconic acid (IA) and lactic acid (LA).

Data is generated, collected and processed during the implementation of the project in order to carry out the tasks described in the VIVALDI Grant Agreement (GA) to reach the specific objectives which were set with the vision of making significant advances when moving towards a new CO₂-based sector that is environmentally and economically competitive with the current fossil-based alternatives. Collecting and sharing of data not only within the consortium but also with relevant stakeholders is essential for reaching this goal. The collected data will allow the VIVALDI project to pave the path towards near zero-waste biorefinery, to contribute to the EU leadership in CO₂ utilisation, to help BIs to survive and flourish while fighting climate change and promoting circular economy and industrial symbiosis, and to provide evidence-based support for future changes in regulation that is restraining a wider CO₂-based deployment.

The purpose for data collection in each Work Package (WP) is outlined below:

WP1: Electrochemical conversion of BI CO₂ emissions

WP1 is developing, upscaling and validating a reliable methodology for electrochemical conversion of raw CO₂ emissions from BI to methanol (MeOH) and formic acid (FA) at biocompatible conditions. Data is collected to demonstrate adequate CO₂ enrichment and purification of the BI emissions for its electrochemical conversion, efficient electrochemical reduction of CO₂ to FA and MeOH, long-term validation of the technology at higher scale and development of downstream solutions to obtain enriched FA/MeOH solutions suitable for posterior bioconversion.

WP2: Nutrient recovery from BI wastewaters

WP2 is developing and validating a bioelectrochemical technology for nutrient recovery from the industrial wastewater to serve as microbial feedstock for a subsequent yeast-based fermentation. Data is collected in order to develop the technology for the recovery of ammonia and macronutrients in microbial electrolysis cells (MECs) and bioelectroconcentration cells (BECs), respectively, with high rate, high coulombic selectivity and lower energy demand than current state-of-the-art (SOTA). The lab-scale results will be validated at a higher scale using a selected real wastewater for its integration with the fermentation step.

WP3: Bioproduction of added-value organic acids

WP3 is developing and validating an integrated efficient yeast-based fermentation process for the conversion of CO₂, MeOH and FA to added value organic acids (3-HP, IA, LA and SA). For this aim, the data previously generated at BOKU and UAB about *Pichia pastoris* metabolism is used to engineer the adequate yeast strains. Data is collected and generated in order to study the fermentation process for each targeted organic acid at lab-scale to obtain high production yields at adequate downstream processing conditions for WP4. Moreover, CO₂ reduction and fermentation steps will be integrated in a single reactor.



WP4: Downstream and industrial validation

WP4 is validating the produced CO₂-based organic acids in the industrial flowchart of BI. Data is collected to design adequate downstream processes to ensure that the fermentation products comply with the specifications required by the end-users. The validation of the VIVALDI concept will be conducted in an industrial relevant environment. VIVALDI considers an iterative process with WP3 to further refine the project results out of the conclusions and recommendations from the industrial partners.

WP5: Sustainability and circularity assessment

WP5 is quantifying and ensuring that VIVALDI's solutions are sustainable through a quadruple bottom approach integrating circularity, environmental, economic and social dimension. Data is being collected to evaluate the environmental positive benefits of the whole VIVALDI loop using a Life Cycle Assessment (LCA) in terms of greenhouse gas (GHG) emissions, ecotoxicity, use of resources and energy demand and a circularity assessment that will identify and quantify opportunities to adapt circular strategies. A Life Cost Cycle (LCC) is being used to compare the cost-effectiveness and competitiveness of VIVALDI versus current SOTA. Finally, a social LCA is being used to evaluate the socio-economic impacts and benefits of VIVALDI's implementation.

WP6: Exploitation and regulatory aspects

WP6 is identifying market opportunities for each of VIVALDI's solutions (i.e. technologies and products) by mapping and engaging key stakeholders and by clearly defining the new business models and exploitation roadmaps for CO₂-producers, end-users and technology adopters. Data is being collected for making VIVALDI an attractive concept for the market and for developing the required exploitation actions for a fast market adoption of VIVALDI's solutions.

WP7: Communication and dissemination

WP7 is raising awareness and efficiently disseminating the knowledge generated in VIVALDI to all the actors of the society from stakeholders to researchers in view of maximising the impact of VIVALDI's outcomes in adequate scientific diffusion channels and of ensuring replicability throughout Europe's industrial hubs. Data is being collected to actively communicate and disseminate VIVALDI's results to the broadest range of audiences possible.

WP8: Management

The main goal of WP8 is to ensure a timely execution of the project and the accomplishment of its objectives. The policies and guidelines for handling all the data collected, processed and generated during VIVALDI are being managed as a part of WP8.

WP9: Ethics requirements

This work package sets out the ethics requirements that the project must comply with. The policies for handling personal data have been included in the data protection policy (D.9.2).



2. Collected and generated data

Due to the comprehensive multi-aspect approach of VIVALDI, the information has been and will be collected from various sources. Information regarding the origin, type, format and size of datasets collected or generated during the implementation of VIVALDI are collected from each beneficiary using a questionnaire (Annex I). The questionnaires will be annually updated to ensure that all the relevant information is included. Types of data collected during the VIVALDI project include:

- Experimental and observational data collected using measurement devices and analytical instruments
- Derived data, derived from the experimental data or literature data
- Survey and interview data, from VIVALDI partners, potential stakeholders and focus groups
- Market-related information
- Methodologies and workflows, standard operating procedures
- Multimedia and physical documents (reports, spreadsheets, presentations) to present VIVALDI data
- Images
- Videos
- Recorded data (interviews)
- Laboratory notebooks, field notebooks, diaries

The main formats of data generated/collected will include:

- Reports, spreadsheets, presentations - Microsoft Office formats, .pdf
- Observed data (experimental) - original format, .csv, .mpr, .mps, .txt, .fasta, .fastq
- Images – .jpg, .gif, .png
- Audio – .mp3
- Videos – .mpg, .avi, .mp4

2.1. Reuse of existing data

As VIVALDI is committed to expand the knowledge beyond the SOTA, previously generated datasets are being utilised to take advantage and build up on the past advances made in the field. The existing data utilised comes mainly from previous data collected by the partners and from scientific literature. In addition, public and commercial databases are being utilised to collect required information. Specific envisioned instances for the re-use of existing data including but not limited to:



- Peer-reviewed published data and data disclosed in public repositories are being studied when implementing the project tasks
- In WP2 and WP3, sample data (e.g., from sequencing) is being compared to public databases
- In WP4, existing data from past research projects with *P. pastoris* is being utilised
- In WP5, sustainability assessment data included in commercial databases is being used to complement the activity data obtained from project activities
- For WP6, publicly available data about the markets is being collected via online search

2.2. Origin of the data

The data will be derived from numerous sources throughout the project, including:

- Lab trials and experiments
- Pilot experiments
- Literature
- External databases
- Stakeholders
- From VIVALDI partners' previous and/or ongoing projects (when possible)
- Data subjects who interact with the VIVALDI website
- Data subjects (surveys, interviews, training workshops, questionnaires)
- Mathematical simulations

2.3. Size of the data

Expected size of data collected in each WP is presented in Table 1. Size values will be revised during project execution.

Table 1: Type and estimated size of data collected in the WPs.

WP	Types of data	Estimated size
WP1	Experimental and analytical data (.xlsx, .csv)	<2 GB
WP2	Reactor designs (.stl, .pdf)	<2 GB
	Experimental and analytical data (.xlsx, .csv) Sequencing data (.fasta, .fastq)	<2 GB <1 TB
WP3	Experimental and analytical data (.xlsx, .csv)	<2 GB



WP	Types of data	Estimated size
	Sequencing data (.fasta, .fastq)	<1 TB
WP4	Experimental and analytical data (.xlsx, .csv, .docx)	<1 GB
WP5	Life Cycle Inventories (.json, .xml, .xlsx)	<500 MB
WP6	Stakeholders Contact Database (.xlsx)	<5 MB
WP7	Website user data and event registrations	<5 MB
WP8	Documentation (.xlsx, .docx, .pdf)	<100 MB
	Sample Lifecycle Management System (.xlsx)	<100 MB
WP9	Documentation (.xlsx, .docx, .pdf)	<100 MB

2.4. Utilisation

The generated data is useful for the VIVALDI consortium partners for their own research and for exploitation of the data generated in the project. Datasets are actively shared within the consortium to ensure successful transfer of knowledge. In addition, the generated data is of interest for researchers outside the consortium working on related fields, including but not limited to:

- CO₂ capture and utilisation
- Treatment of flue gas
- Electrochemical systems
- Wastewater treatment
- Bioelectrochemical systems
- Nutrient recovery
- Biotransformations
- Sustainability

The generated data is also of significant interest to people working in BIs that desire to turn their CO₂ emissions into high value products. The generated data is also raising awareness to all societal actors (general audience, stakeholders, researchers and policymakers) on the solutions that VIVALDI can provide as well as on the need of a CO₂-based industrial sector to fight climate change as well as boosting EU's economy.



3. FAIR Data

VIVALDI is committed to follow the FAIR principles, aiming to make the collected data Findable, Openly Accessible, Interoperable, and Re-useable.

3.1. Making data findable

3.1.1. Naming convention

To ensure that the data is findable, all the data objects are being made uniquely and persistently identifiable. A consistent naming convention is used to link the documents to other data in the project. The naming convention to be used for documents in the VIVALDI project is as follows:

VIVALDI_WPX_L_N_YYYYMMDD_vX.Y	
WPX	Work Package This can be left out for documents that are not related to a specific WP.
L	Label and number of the dataset/document Examples: Deliverable 8.3 – D8.3 Technical report M24 – TR_M24 Financial report M24 – FR_M24
N	Name of the dataset/document
YYYYMMDD	Date
vX.Y	Version



Version numbers incremented at each revision are included in the filenames. The version control is as follows:

- Initial drafts (before submission) are labelled as v0.X
- Revisions to draft prior to submission are numbered in ascending order (e.g. v0.1, v0.2, v0.3)
- The submitted document (final draft) are labelled as v1.0
- If a document needs to be resubmitted, or requires major changes, the first number in the version identifier increases by 1 (e.g. v2.0, v3.0)

3.1.2. Metadata

A data object should be accompanied with metadata that makes it easily findable for both humans and computers. Descriptive metadata is essential also for allowing stakeholders outside the project to re-utilise the generated datasets. The metadata needs to include information on how the data were generated and under what license and how they can be re-used. This metadata is being provided separately from the original raw research data.

The relevant metadata and its presentation depend on the type of the dataset and the used repository. Therefore, the metadata standard used is being individually selected for each dataset. For datasets submitted to CORA, the template included in Annex II is used to collect the appropriate metadata. The template or a modified version of it is utilised also with other repositories, if there no other specific metadata templates are preferred. If no specific recommended standard exists for the dataset type or the repository, DataCite's metadata standard¹ will be followed.

Examples of metadata being included:

- Title of project (including 'European Union (EU)' 'Horizon 2020', project acronym and grant number)
- Title of dataset
- Name of data owner
- Date of submission
- Subject
- Description of data
- Format of data
- Type of data (e.g., data set, article, audio file, image)
- Digital Object Identifier (DOI)
- Access rights
- Keywords

¹ DataCite Metadata Working Group, 2021. DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs. Version 4.4. Available at: <https://doi.org/10.14454/3w3z-sa82>



To ensure that the data is easily findable, the metadata of all datasets deposited in repositories are including keywords, which are selected from controlled vocabularies that are suitable for the specific type of data, (e.g., following the officially recommended analytical nomenclature mandated by IUPAC).

The bibliographic metadata includes the following:

- the terms “European Union (EU)” and “Horizon 2020”,
- the name of the action, acronym and grant number,
- the publication date and the length of embargo period, if applicable,
- a persistent identifier

3.2. Making data available

Data sets are shared internally with the consortium using the Teams platform. For open sharing, UAB’s proprietary Research data Management Systems is being adopted to facilitate the storage of subsets of data (technical fact sheets, metadata of experimental datasets and their computational analysis). Public documents (e.g., deliverables, research articles, datasets and protocols) are shared using the digital document depository of UAB (<http://ddd.uab.cat/>). In addition, datasets are being shared publicly also using CORA, a Catalan multidisciplinary data repository for publishing research data sets in FAIR mode (<https://dataverse.csuc.cat/>). Both systems provide unique URLs to access the files. Alternatively, the beneficiaries can choose to share their data openly in institutional or public repositories. One public option for open sharing is Zenodo (<https://zenodo.org/>), an open online research data repository, the repository structure, facilities and management of which follow the FAIR data principles.

Data is being made accessible through online repositories and, therefore, specialist software tools should not be required to access the data. Whenever possible, the data is shared in a format that is readable with cost-free software. If special software or tools are required for accessing the data, a short text file (e.g., ASCII) is provided with the data file to explain the software required. The research data and associated metadata, documentation and code are being deposited to one of the aforementioned repositories.

Potential restrictions to sharing datasets as open access are being reviewed separately for each dataset. Access to confidential or restricted data will be awarded under the discretion of the Work Package Leader (WPL) of the WP that generates the data, alongside the coordinator. VIVALDI does not currently have a data access committee, but the need for such committee will be again evaluated as the project proceeds and the type and the amount of data generated sharing of which is restricted is clearer.

To protect the copyright of the knowledge generated in the project, Creative Commons license will be used in certain circumstances. The access to the data repository will be layered, integrating a pertinent licensing system, with the purpose of limiting the accessibility to the different types of data depending on the data owners’ agreement. Any associated metadata, documentation and code will also be deposited in the repository. The owner of the data will determine which of these licenses will be used when data is posted on the repositories. Zenodo provides well-described conditions for access (see <http://about.zenodo.org/policies/>).



When data is publicly shared, the person accessing the data is not be directly identified. The users are expected to follow 'The European Code of Conduct for Research Integrity'².

3.2.1. Making data openly accessible

VIVALDI participates in the Open Research Data Pilot and thus follows the EU guidelines described in the document 'Guidelines to the Rules on Open Access to Scientific Publications and Research data in Horizon 2020'³. The data is being made available following the 'as open as possible, as closed as necessary' principle, meaning that all the data is being made available for verification and re-use, unless the owner of the data can justify why data cannot be made openly accessible. To maximise the impact of the VIVALDI project, selected data and results are shared with the scientific community and other stakeholders by publishing them in scientific journals, by presenting them at scientific conferences and by sharing them in open access data repositories. Unless there is a specific reason why the data cannot be made public, all the data associated with scientific publications is being made openly available. The reasons why data cannot be made publicly available include, but are not limited to:

- Sharing is restricted for legal and contractual reasons
- Data was obtained with the permission of third parties, but the third parties have not agreed to make the data publicly available
- Data compromises the protection of a partner(s) intellectual property
- Data is commercially sensitive

VIVALDI follows an open access policy by providing on-line access to scientific information that is free of charge to the end-user and that is re-usable. In the context of this project, scientific information refers to peer reviewed scientific research articles (published in scholarly journals), pre-print articles, conference papers, patents, books and research data (data underlying publications, curated data and/or raw data). Each consortium partner commits to deposit each latest publication as soon as possible. Each partner will ensure open access to the publications immediately, if an electronic version is available for free via the publisher, or within six (6) months of publication, if it is not.

3.3. Making data interoperable

To allow for interdisciplinary interoperability, the data includes standard vocabularies for all data types present in the data sets where possible. In the case that the usage of uncommon or project specific ontologies/vocabularies is unavoidable, mappings to more commonly used ontologies will be provided. To ensure that the data can be easily read and understood, it is accompanied by standardised metadata. The relevant metadata and its presentation depend on the type of the dataset and the used repository and thus the metadata standard used is being individually selected for each dataset.

² The European Code of Conduct for Research Integrity, 2017. All European Academies (ALLEA). Available at: <https://allea.org/code-of-conduct/>

³ Guidelines to the Rules on Open Access to Scientific Publications and Research data in Horizon 2020, 2017. European Commission. Version 3.4. Available at: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf



3.4. Increasing data reuse

VIVALDI commits to sharing data as openly as possible. However, to protect the ownership of the datasets, Creative Commons licenses will be assigned in certain circumstances. The owner of the data will determine which of these licenses will be used when data is posted on the repositories. For the datasets that are publicly shared, ShareAlike and NonCommercial-ShareAlike licenses can be considered. For data used in articles published in open access journals, data sharing may be limited due to an embargo period. European Commission recommends a maximum embargo period of six (6) months. Access to datasets deposited on public data repositories such as Zenodo will be unlimited.

Restrictions on re-use policy will be applied on data that requires protection and can be re-used only by the consortium. The following restrictions may be applied with Creative Commons licensing of the dataset:

- Attribution: requires users of the dataset to give appropriate credit, provide a link to the license, and indicate if changes were made.
- NonCommercial-ShareAlike: prohibits the use of the dataset for commercial purposes by others.
- ShareAlike: requires the others to use the same license as the original on all derivative works based on the original data.

Datasets will not be published before the date of publication of the associated scientific article, considering any restrictions and embargo periods. Data Reports will become available after each related deliverable if the dissemination level allows it. Data produced and/or used in the project will be useable by third parties for any purpose dependent on the selected licences and confidentiality requirements. The data that is re-useable will be intended to be so for its lifetime.

Files classified as public will be accessible and downloadable for at least five (5) years after the end of the project. Files classified as internal and/or confidential will be accessible only to project partners from the internal platform for at least five (5) years after the end of the project. If Zenodo is used as the online repository, any data stored in the repository will be retained for the entire existence of the repository, as is the policy of Zenodo.

4. Quality assurance

To ensure the quality of any experimental data, all personnel involved in testing are properly trained for data quality management. Data is being analysed by the partner generating/collecting the data in order to identify any issues such as missing values, anomalies, or inconsistencies. Each experiment is being characterised in terms of accuracy and precision of the significant data measured during the test. Inconsistent, inaccurate and/or missing data will be managed by repeating the experiment. No data without information on consistency will be considered for technical purposes and/or modelling.

Templates are being used for project documentation, and an internal peer review is performed for the main project deliverables to guarantee that the deliverables are developed with a high level of quality. Each lead beneficiary has to submit the documents to the partner assigned as internal reviewer as well as to the project coordinator for review.



5. Costs

Costs related to open access of the research data in Horizon 2020 are eligible for reimbursement under the conditions defined in the H2020 GA. Project beneficiaries are responsible for applying for reimbursement for costs related to making data accessible to others beyond the consortium. If a cost-free repository is used, there are no allocation of resources necessary for data storage.

The project coordinator and the innovation board are responsible for the overall data management in the project, with the help of the WPLs. However, each partner has to respect the policies set out in this DMP and the datasets have to be created, managed and stored appropriately and in line with the applicable legislation. The partner that generates the data is responsible for the validation and registration of the datasets, for providing the required metadata and for sharing the data through the open access repositories.

If the selected repositories do not provide the desired storage length, additional resources for long-term preservation of data will be evaluated based on the type and the size of the data requiring long-term preservation.

6. Data security

Data from the project is being safely stored in certified repositories for long-term preservation and curation with General Data Protection Regulation 2016/679 (GDPR 2016/679) compliant features and security measures. The server ensures that the regulations and security procedures are compiled following the EU standards. Archiving and preservation of data generated by the VIVALDI partners at their respective institutions is the responsibility of the partners individually.

To ensure the security of sensitive data, the following guidelines are followed:

- Files are named systematically to ensure coherence of the final dataset
- Data is stored in at least two separate locations to avoid the loss of data
- Data is encrypted, if necessary
- The use of Universal Serial Bus (USB) flash drives will be minimised

All project deliverables and data are being stored and shared in the UAB's Teams folder restricted to the project consortium. Initially, only the consortium members have access to the folders where datasets and metadata are filed. Security of data shared on the additional internal platforms will be ensured by requiring login details.

7. Ethical aspects

In addition to the European Commission policies on open data management, project partners must also adhere to their own institutional policies and procedures for data management. The WP9 aims at ensuring that ethical requirements are met for all research undertaken in the project, including data management



aspects, in compliance with H2020 ethical standards. All partners will assure that the EU standards regarding ethics and data management are fulfilled in compliance with the ethical principles (see GA, Article 34) and confidentiality (see GA, Article 36). In addition, in accordance with the GDPR 2016/679, the data controllers and processors are fully accountable for the data processing operations. The ethical considerations of data collection in the VIVALDI project are detailed in the deliverables D9.1 and D9.2. These will be strictly adhered by the VIVALDI partners. The deliverable D9.1 also contains templates for informed consent forms and information sheets.

8. KPI Management Plan

The progress towards the VIVALDI's goals is being tracked using Key Performance Indicators (KPIs). Currently, 52 KPIs have been identified and additional KPIs will be added, if deemed beneficial for evaluating the progress towards achieving the set objectives. The progress in each KPI is being followed monthly using the KPI management system in Teams. Each WPLr in collaboration with the Task Leaders is responsible for updating the progress of the KPIs monthly. This monthly reporting allows the WPLs and the project coordinator to continuously monitor the progress achieved with all the KPIs and to react promptly to deviations or delays versus the anticipated/expected progress. In addition to the monthly monitoring, KPI's are annually reviewed to evaluate their validity and the need for new KPIs.

As soon as delays or deviations are detected, the reasons will be identified together with the beneficiaries participating in the related tasks. Corrective actions taken can include but are not limited to 1) re-assessing the technical approach, 2) assigning additional resources to the task, 3) consulting external advisors and 4) re-assessing the KPI's target value.

9. Additional resources

This DMP has been created using the tool "EinaDMP" (<https://dmp.csuc.cat>).

10. Annexes



ANNEX I: Questionnaire regarding collected data

VIVALDI Data Management

Please complete the following questionnaire regarding the datasets that are generated and used during the implementation of the VIVALDI project and the handling of data according to your current knowledge.

Partner:	
-----------------	--

Are you handling personal data?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Do you have appointed Data Protection Officer?	<input type="checkbox"/> Yes	<input type="checkbox"/> No

Institutional policies and procedures for data management:
Please describe the policies and procedures that your institution has in place that can affect the handling and sharing of the data that is generated and/or used in the VIVALDI project. If these available as a document that you can share, please return them with this questionnaire.

Are you planning to use any institutional and national data repositories for sharing the data of VIVALDI project?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Please provide a description (and a link, if available) of the relevant repositories. You may include information also about repositories that you think that could be of interest to other VIVALDI participants.		



Datasets generated during the project:

Fill in the following table for each **dataset expected to be generated** or an **existing dataset** that is used in the VIVALDI project. Fill in as many of the fields as possible, but if the information is not available yet, the cells can be left blank and updated later.

Dataset

Item	Description
Dataset identifier (WP_Task_partner_number, e.g. WP2_T2.1_UAB_1)	
Dataset name	
Dataset description	
Contains personal data	<input type="checkbox"/> Yes <input type="checkbox"/> No
Informed consent required	<input type="checkbox"/> Yes <input type="checkbox"/> No
Key contact [Beneficiary]	
Other contributing beneficiaries	
Dataset file format and size (estimation, if not know)	
Other associated tasks (if any)	
Keywords	
Metadata standard, if specific	



<p>To whom this data is useful for? (e.g. other researcher working on (which?) fields, policy makers, public and industrial stakeholders)</p>	
<p>Key data sources/Origin of the data</p>	
<p>Are there any ethical or legal issues that can have an impact on data sharing?</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No
<p>If yes, please elaborate</p>	
<p>Dataset can be made public</p>	<input type="checkbox"/> Yes <input type="checkbox"/> No
<p>If No, justification:</p>	
<p>Repository(-ies) (if known)</p>	
<p>Dataset DOI</p>	
<p>Dataset version history</p>	



ANNEX II: CORA Metadata template

[Instructions in this document are in between brackets.]
 [When you are done filling this template delete all instructions and delete any sections or questions that do not apply to the dataset.]
 [Fill this template as thoroughly as possible to ensure the reusability of your dataset.]
 [This template was created by the Library Service of the Universitat Autònoma de Barcelona from templates prepared by: Research Data Services at Oregon State University, Research Data Management Group - Cornell University, TUDelft Library and University of Bath Library.]
 [For questions or guidance about using this template contact ddd.bib@uab.cat]
 [This template is published under a CC0 license.]

This documentation file was generated on [date in YYYY-MM-DD format] by [Name]

GENERAL INFORMATION

1. Title of Dataset:

2. Authorship:

[Fill out the information about researchers according to this structure. Repeat the process, one for each author.]

Name:
 Institution:
 Email:
 ORCID:

DESCRIPTION

1. Dataset language:

2. Abstract:

[The abstract should describe the dataset, not the research or the results obtained after analyzing the dataset. The dataset abstract should be different from the article, book or thesis abstract. The abstract can be written in several languages.]

3. Keywords:

[A minimum of three keywords. You can write them in different languages.]

4. Date of data collection (single date or date range):

[Format YYYY-MM-DD or YYYY-MM-DD - YYYY-MM-DD]

[Do not copy paste the methods section from a pending publication unless you have made sure that you can do that: some journals may consider this as a publication, and will not accept the manuscript.]
 [If you want to refer to an article that has not been accepted for publication yet, include as much information as you have at the moment.]
 [Include the DOI in the citation.]

1. Description of methods used for collection-generation of data:

[Include references, in your discipline's standard style, to publications or other documentation containing experimental design or protocols used in data collection.]

2. Methods for processing the data:

[Describe how the submitted data were generated from the raw or collected data.]

3. Instrument- or software- specific information needed to interpret the data:

[Include version of software. If software is not openly available, explain where to get the software. Consider changing the format of the dataset into an open format (recommended).]

4. Instruments, calibration and standards information:

5. Environmental or experimental conditions:

[E.g.: atmospheric influences, computational environment, etc.]

6. Quality-assurance procedures performed on the data:

FILE OVERVIEW

[All files in the dataset should be listed here. Filenames should include the extension (.csv, .pdf, etc.)]

1. Explain the file naming convention, if applicable:

2. File List:

[Repeat the structure for every file. Include directory/folders structure, if applicable.]

Filename:
 Short description:

5. Date of dataset publication:

[Format YYYY-MM-DD]

6. Funding sources:

[Repeat the information for each funder if applicable]

Funding agency:

Project number:

7. Geographic location/s of data collection:

[Latitude, longitude, city, region, country, etc.]

ACCESS INFORMATION

1. Creative Commons License of the dataset:

[Recommended license: CC0. See all the CC licenses recommended by the UAB: <http://ddd.uab.cat/record/129285>]

2. Dataset DOI:

[This information will be filled in by the DDD team after validating data.]

3. Related publication:

[Bibliographic citation of the publication related with the dataset in your discipline's standard style, including DOI.]

4. Link to related datasets:

[(E.g. Datasets from the same project) according to this structure.]

DOI/URL:

VERSIONING AND PROVENANCE

1. Last modification date:

[Format YYYY-MM-DD]

2. Was data derived from another source?:

[Answer Yes or No. If Yes, list source in your discipline's standard style.]

3. Additional related data collected that was not included in the current data package:

METHODOLOGICAL INFORMATION

[Describe the methodology used to generate the dataset.]

3. Relationship between files:

4. File format:

TABULAR DATA-SPECIFIC INFORMATION

This section should be created for each file or dataset that requires explanation of variables (typically, tabular data). All variables should be described. Include the measurement units.]

1. File name:

2. Number of rows and columns:

3. Variable list:

[Repeat the structure for each variable.]

Variable name:

Description:
 Measurement units or value labels:

4. Missing data codes:

[Repeat the structure for each code or symbol.]

Code or symbol:

Defintion:

5. Specialized formats or abbreviations used:

MORE INFORMATION

[Include any other information about the dataset not reflected in the template that you consider relevant]